PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | |
|---|---|---|
| (51) International Patent Classification 6 : C07K 14/435, C12N 1/21, 5/10, 15/09, 15/12, 15/53, 15/63, C12P 17/18, C07D 471/00, C12Q 1/68 | A1 | (11) International Publication Number: WO 95/21191<br>(43) International Publication Date: 10 August 1995 (10.08.95) |

(21) International Application Number: PCT/US95/01425

(22) International Filing Date: 3 February 1995 (03.02.95)

(30) Priority Data:
08/192,158 4 February 1994 (04.02.94) US

(71)(72) Applicants and Inventors: WARD, William [US/US]; 149 E. Chestnut, Metuchen, NJ 08840 (US). CHALFIE, Martin [US/US]; 15 Claremont Avenue No. 51, New York, NY 10027 (US).

(74) Agent: BEHR, Omri, M.; 325 Pierson Avenue, Edison, NJ 08837 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: BIOLUMINESCENT INDICATOR BASED UPON THE EXPRESSION OF A GENE FOR A MODIFIED GREEN-FLUORESCENT PROTEIN

(57) Abstract

A pre-coelenterazine peptide comprising a modified *A. victoria* GFP having an amino acid sequence in which $Ser^{65}$ is replaced with Tyr. There are further provided a polynucleotide encoding the pre-coelenterazine peptide, allowing synthesis of large, pure amounts of coelenterazine, as by culturing organisms transformed with the polynucleotide; methods for synthesizing coelenterazine; and improved assays employing the polynucleotide or transformed organisms, e.g., to detect mutagenesis.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | GB | United Kingdom | MR | Mauritania |
| AU | Australia | GE | Georgia | MW | Malawi |
| BB | Barbados | GN | Guinea | NE | Niger |
| BE | Belgium | GR | Greece | NL | Netherlands |
| BF | Burkina Faso | HU | Hungary | NO | Norway |
| BG | Bulgaria | IE | Ireland | NZ | New Zealand |
| BJ | Benin | IT | Italy | PL | Poland |
| BR | Brazil | JP | Japan | PT | Portugal |
| BY | Belarus | KE | Kenya | RO | Romania |
| CA | Canada | KG | Kyrgystan | RU | Russian Federation |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SD | Sudan |
| CG | Congo | | | SE | Sweden |
| CH | Switzerland | KR | Republic of Korea | SI | Slovenia |
| CI | Côte d'Ivoire | KZ | Kazakhstan | SK | Slovakia |
| CM | Cameroon | LI | Liechtenstein | SN | Senegal |
| CN | China | LK | Sri Lanka | TD | Chad |
| CS | Czechoslovakia | LU | Luxembourg | TG | Togo |
| CZ | Czech Republic | LV | Latvia | TJ | Tajikistan |
| DE | Germany | MC | Monaco | TT | Trinidad and Tobago |
| DK | Denmark | MD | Republic of Moldova | UA | Ukraine |
| ES | Spain | MG | Madagascar | US | United States of America |
| FI | Finland | ML | Mali | UZ | Uzbekistan |
| FR | France | MN | Mongolia | VN | Viet Nam |
| GA | Gabon | | | | |

# BIOLUMINESCENT INDICATOR BASED UPON THE EXPRESSION OF A GENE FOR A MODIFIED GREEN-FLUORESCENT PROTEIN

## Field of the Invention

This invention concerns a pre-peptide of coelenterazine which undergoes transformation to coelenterazine; a polynucleotide which encodes for the pre-peptide of coelenterazine; living organisms transformed with this polynucleotide; methods for synthesizing coelenterazine; and improved assays employing the polynucleotide or transformed organisms, *e.g.,* to detect mutagenesis.

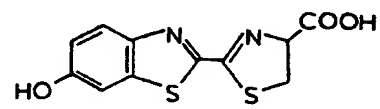
## BACKGROUND OF THE INVENTION

For some years, it has been appreciated that bioluminescence offers a useful indicative tool in a variety of assays. Assays employing bio-luminescence enjoy the advantages of accuracy and great sensitivity. The accuracy results from the highly specific interaction between luciferase and its luciferin; consequently, the number of false positive indications is minimized. The sensitivity is due to the great sensitivity of light sensing equipment and photomultipliers. When properly designed, such assays may additionally offer a quantitative relationship between the level of light released and the phenomenon being measured.

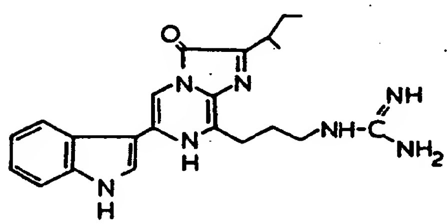
The general chemical reaction underlying the phenomenon of bio-luminescence is the oxidation of a substrate ("luciferin") by an enzyme ("luciferase"), usually in the presence of oxygen. An intermediate, ener-gized "oxyluciferin*," is formed during the oxidation reaction, which, in proceeding to the oxidized form ("oxyluciferin"), releases light. This inter-action between luciferin and luciferase is seen in the anthazoan coelenter-ates such as the sea pansy *Renilla reniformis*.

It should be noted that the terms "luciferin" and "luciferase" are non-specific: they are used in the literature to refer to the bioluminescing substrate and enzyme of nearly all bioluminescing organisms. The specific luciferins found in nature however vary extensively. For example, the luciferin of the firefly, the ostracod *Cypridina* and the coelenterates (a group of marine organisms including the jellyfish and the sea pansy) are clearly distinct molecules, as shown in Formulae I, II and III respectively:

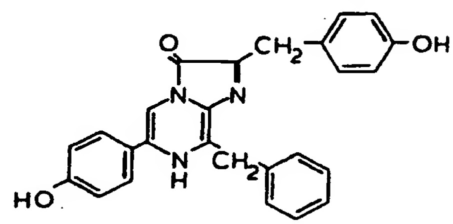
Formula I



Formula II



Formula III



These three compounds are usually termed firefly luciferin, *Cypridina* luciferin and coelenterate-type luciferin (or "coelenterazine") respectively to avoid confusion.

Generally, there is little cross reaction between the luciferase of one species and the luciferin of another: firefly luciferase will not oxidize either the *Cypridina* or coelenterate-type luciferin. Among members of an animal phylum however, (e.g., among coelenterates), cross reactions are found: thus the combination of *Aequorea victoria* coelenterazine with *R. reniformis* luciferase does generate bioluminescence. This occurs because all coelenterates have coelenterazine as their luciferin.

Recombinant DNA techniques have helped to elucidate bioluminescent systems and thus have enabled investigators to introduce some bioluminescent compounds into assays to provide a convenient, conspicuous indicator. For example, the firefly luciferase gene has been expressed in *Escherichia coli*, tuberculosis bacilli and tobacco plants grown in an ambient medium containing firefly luciferin. Also, genes for the bacterial *lux* system have been introduced into a number of non-bioluminescent species. These bacteria transformed with *lux* genes have been employed in a wide range of assays discussed in Gould and Subramani Anal.Bioch 175, 5-13; Stewart and Williams, J.Gen.Microbiol., 138 1289-1300 (1992); Stewart and Williams, ASM News, 59, No. 5, 241-246 (1993); and Hill et al., Biotechn. App.Bioch. 17, 3-14 (1993); all incorporated herein by reference.

These assays are limited however by the requirement of most bioluminescent systems that luciferin, a compatible luciferase and an exogenously added co-factor all be present. In the absence of the substrate, enzyme or co-factor, the system does not bioluminesce. Thus, firefly and bacterial luciferases will only oxidize firefly luciferin and $FMNH_2$ if ATP or organic aldehydes are present respectively. And since most isolated genes to date have been for enzymatic luciferases -- the biosynthetic pathway of luciferins remaining largely unclear -- it has been necessary to grow organisms transformed with a luciferase gene in media containing the luciferin.

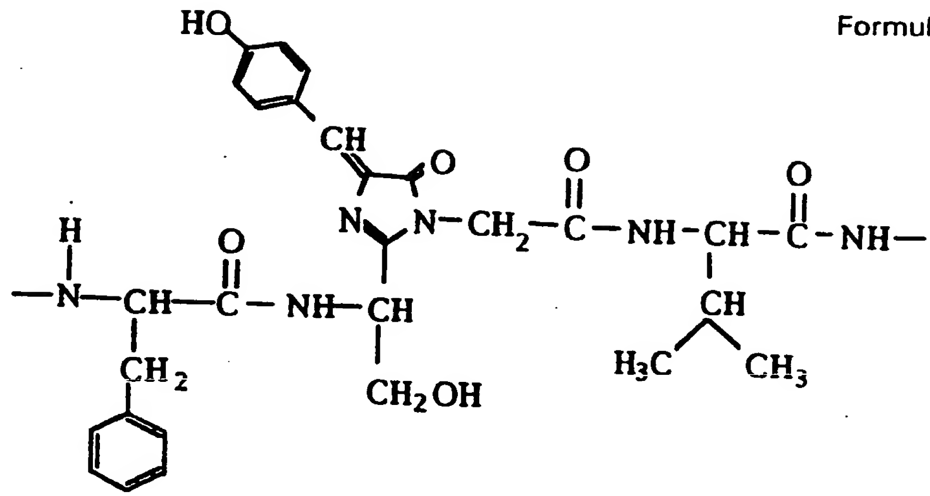
This requirement that both luciferase and luciferin be present has limited non-bacterial bioluminescing compounds to assays where the cells under investigation are transformed with a luciferase and either have cell walls and plasma membranes which are permeable to a compatible luciferin or which are rendered permeable at some point. Thus, in Ow et al., Science 234, 856-859 (1986), tobacco plants grown from cells transformed with the firefly luciferase gene were exposed to a liquid medium containing firefly luciferin. The plants exhibited bioluminescence primarily along their major

veins. Alternatively, lysis of *E. coli* transformed with a firefly luciferase gene was employed by Lee et al., Anal.Chem. 64, 1755-1759 (1992), incorporated herein by reference, to indicate mutagenesis. Upon induction of a lysogenic bacteriophage carried by the *E. coli*, bacterial lysis released the firefly luciferase into a medium containing firefly luciferin. The level of bioluminescence in the bacterial growth medium indicated the level of mutagenesis.

Because these assay systems relied upon cell walls or membranes which were permeable or had been made permeable to luciferin, they were limited to use either on living organisms having a particular cell wall or membrane permeability, or on lysed cells. Recently, however, the need for cell wall or membrane permeability or for cell lysis has been obviated in US Patent application Serial No. 08/119,678, filed September 10, 1993 by Chalfie et al. (hereinafter "Chalfie et al.") This application, which is incorporated herein by reference, describes the synthesis in *E. coli* and *Caenorhabditis elegans* of the "green-fluorescent protein" (hereinafter "GFP") of the jellyfish *A. victoria*.

GFP is a polypeptide derived from an apopeptide having 238 amino acid residues and a molecular weight of approximately 27,000. GFP contains a chromophore formed from amino acid residues 65 through 67. Investigators have proposed a mechanism for the formation of the GFP chromophore. In this proposed mechanism, $Tyr^{66}$ in GFP is dehydrogenated, and later cyclizes along with its upstream neighbor $Ser^{65}$, as well as its downstream neighbor $Gly^{67}$ to form the imidazole ring chromophore having the structure shown in Formula IV.

Formula IV



Cody et al., <u>Bioch</u>. <u>32</u> 1212-1218 (1993).

As its name indicates, GFP fluoresces; it does not bioluminesce. *In vivo*, the chromophore of GFP seems to be activated by energy transfer from coelenterazine complexed with the photoprotein aequorin, as in the hydrozoa order of coelenterates such as the jellyfish *A. victoria*. Organisms containing GFP thus exhibit green fluorescence at 510 nm, rather than the blue wavelength light at 480 nm typical of coelenterazine bioluminescence.

In the system of Chalfie et al., cells were transformed with a cDNA for the 238 amino acid apo-GFP. Expression of this apoprotein of GFP in the absence of other jellyfish gene products and was said to result in post-translational modification at residues 64 through 69 to form the GFP chromophore (of Formula IV above). The resulting GFP was said to exhibit the characteristic green fluorescence at 510 nm upon irradiation with blue or UV light. Thus, cells transformed with the cDNA for apo-GFP may be tested for GFP expression simply by irradiation with blue or UV light. No cell lysis is required to detect fluorescence; one need not provide a co-factor or a luciferin.

Unlike most bioluminescent systems, which require one or more co-factors or a luciferase enzyme in order to release light, GFP fluoresces when illuminated with certain wavelengths of light. Thus, organisms transformed with the GFP gene can exhibit fluorescence while alive and without lysis, unlike the firefly luciferin of Ow et al., *supra*. Further, the GFP gene may be operatively linked with a duplicate of a promoter controlling expression of a protein of interest; expression of this protein can therefore be monitored in a living cell by the detection of GFP's fluorescence. Moreover, transformation of cells of transparent organisms, such as *C. elegans* or zebra fish, permits one to determine the progeny of these cells as the organism develops. The GFP system thus provides a tool for detecting specific physiological events *in vivo* as well as for tracking expression of proteins of interest.

Coelenterazine is a bioluminescing compound commonly found in organisms which synthesize GFP. For example, the jellyfish *A. victoria* produces both compounds. Unlike GFP, coelenterazine is a small non-proteinaceous, highly complex molecule. Coelenterazine (or 3,7-dihydro-2-methyl-6-(*p*-hydroxyphenyl)-8-benzylimidazo[1,2-a]pyrazin-3-one, Hori and Cormier, Proc.Nat.Acad.Sci. 70, No. 1, 120-123 (1973) releases blue light across a broad range peaking at 480 nm upon oxidation by luciferase *in vitro*.

Coelenterazine exhibits bioluminescence at 480 nm in the presence of a compatible luciferase and oxygen; it does not require a co-factor. Despite this advantage, coelenterazine has not been widely adopted for use in assays, primarily due to the difficulty and expense of isolating significant amounts of coelenterazine. The compound and its luciferase are present in bioluminescing organisms at exceedingly low levels: forty thousand sea pansies (*R. reniformis*) are required to collect 0.5 mg of coelenterazine, and six thousand sea pansies for a few mg of *Renilla* luciferase. "General Aspects of Bioluminescence," Ward, 321-358, at 344 in Chemi- and Bio-

luminescence, J.G. Burr, Ed., Marcel Dekker, Inc., 1985. Furthermore, these compounds are chemically unstable, having half lives in aqueous systems of one to two hours. While synthetic coelenterazine is available commercially (from London Diagnostics, Eden Prairie, Minnesota), the numerous complex organic reactions involved in its synthesis render it quite expensive. Recombinant DNA techniques to date have provided no assistance, since the natural route of coelenterazine biosynthesis, and indeed that of most luciferins, is unknown.

Some investigators have theorized biosynthetic pathways for coelenterazine. McCapra and Perring, 359-386, at 371, in Chemi- and Bioluminescence, J. G. Burr, *supra*, noted some structural similarities between coelenterazine and the tripeptide tyrosyl-tyrosyl-phenylalanine, and expressed a belief that coelenterazine is derivable from this tripeptide. Elsewhere, McCapra alone pointed out structural similarity between *Cypridina* luciferin (which shares the fused imidazopyrazine ring of coelenterazine) and the tripeptide tryptophanyl-isoleucyl-arginine, and synthesized the *Cypridina* luciferin from a dehydrotripeptide in JCS Chem.Comm. 1972, "Cyclisation of a Dehydropeptide Derivative: a Model for Cypridina Luciferin Biosynthesis" 894-895. Observing a possible connection between a tripeptide and luciferin, in FEBS Lett. 104, 1979, pp. 220-222, Shimomura stated: "Partial similarity between structure B [the proposed GFP chromophore] and the structure of coelenterazine may suggest a biogenetic significance." Ward similarly observed in Chemi- and Bioluminescence J.G. Burr, *supra*, at p. 329 that in view of the chemical similarities between *Cypridina* luciferin and coelenterazine, it was intriguing to speculate that coelenterate-type luciferin and the GFP chromophore may be produced by a common biosynthetic mechanism involving post-translational protein modification (i.e., ring formation) followed by excision of the chromophore in the case of luciferin synthesis.

Despite the references to tripeptides in the biosynthesis of coelenterazine, it has hitherto generally been believed by investigators that coelenterazine was not generated through ribosomal peptide synthesis. McCapra and Perring's observations above were directed to possible laboratory synthesis routes for the luciferins discussed; for they flatly state that "[n]one of the chemical syntheses [of luciferins] are amino acid based...." McCapra and Perring, *supra*, at 371. Rather, due to the compound's structural complexity, investigators expected coelenterazine was synthesized nonribosomally, in the manner of $\gamma$-glutamyl-cysteinylglycine ("glutathione", the natural antioxidant which forms 5-oxoproline in the $\gamma$-glutamyl cycle) or gramicidin. Biochemistry, Voet and Voet, John Wiley & Sons, pp. 709-711, 941-942; and Biochemistry, Zubay, 2d Ed., Macmillan Publishing Company, p. 796. Accordingly, it has been hypothesized that different organisms might possess a series of cooperating enzymes, for the synthesis of coelenterazine, J.W. Hastings, J.Mol.Evol. 19, 309-321 (1983); and that coelenterazine synthesis occurs *in vivo* via a sequence of enzymatic or chemical reactions, McCapra and Perring, *supra*, at 375-376.

The present invention resolves several difficulties in applications of a coelenterazine bioluminescent system.

## SUMMARY OF THE INVENTION

Applicants have discovered that by modifying the cDNA for the apopeptide of *A. victoria* GFP (described in U.S. Patent application Serial No. 08/119,678, filed September 10, 1993), a heretofore unknown pre-coelenterazine peptide is synthesized. Polynucleotides encoding this pre-peptide allow synthesis of large, pure amounts of coelenterazine and enable numerous methods for imparting bioluminescence to organisms under a variety of conditions.

The discovery that GFP may by slight modification be altered to a pre-coelenterazine peptide, and that polynucleotides encoding the apopeptide

for GFP could be modified to encode for the pre-coelenterazine peptide was wholly unforeseen. There existed no earlier indication that coelenterazine results from post-translational modification of any peptide; or that coelenterazine is or could be genetically encoded. Nor was there any clear indication that GFP and coelenterazine might share a common or related precursor.

Moreover, the newly discovered steps believed to be involved in coelenterazine formation from the pre-coelenterazine peptide are substantially different from those believed to take place in the post-translational modification of apo-GFP into GFP. Consequently, the steps involved converting apo-GFP into GFP did not foreshadow the dehydrogenation and cyclization believed to occur in the conversion of the pre-coelenterazine peptide to coelenterazine.

Without in any way limiting the invention, Applicants believe that pre-coelenterazine is transformed into coelenterazine by the dehydrogenation of $Tyr^{66}$ and/or its upstream neighbor $Tyr^{65}$ (which replaces the $Ser^{65}$ of GFP). Either one Tyr residue and one dehydroTyr residue (or both dehydroTyr residues) then cyclize with their further upstream neighbor $Phe^{64}$. In this cyclization, the peptide bond between residues 63 and 64 and that between residues 66 and 67 are broken; cyclization thus brings about excision of coelenterazine from the peptide.

Nothing in the Cody et al. mechanism of GFP chromophore formation indicates that the replacement of $Ser^{65}$ with $Tyr^{65}$ in the modified GFP would lead to dehydrogenation of both $Tyr^{66}$ and $Tyr^{65}$. Nor is there any indication that this replacement would lead to an upstream shift of the residues which cyclize into a ring; or in two fused rings being formed in place of one: in GFP, residues 65, 66 and 67 cyclizing into a pyrazole ring, while in pre-coelenterazine, residues 64, 65 and 66 cyclize into a fused imadazopyrazine ring.

The following improvements are a result of Applicants' surprising discovery. <u>The first embodiment</u> of the invention is a pre-coelenterazine peptide comprising a modified *A. victoria* GFP in which $R^{65}$ is Tyr. Certain of these peptides comprise at least amino acid residues $R^{1}$ through $R^{228}$ of the modified *A. victoria* GFP. In the pre-coelenterazine peptides, several amino acid residues may vary; $R^{80}$ may be Gln or Arg, $R^{100}$ may be Phe or Tyr, $R^{108}$ may be Thr or Ser, $R^{141}$ may be Leu or Met, $R^{172}$ may be Glu or Lys, and $R^{219}$ may be Val or Ile. Regardless of which of the above-identified amino acyl residues is present at residue 80, 100, 108, 141, 172 and 219, the resulting pre-coelenterazine peptide is formed, when generated *in vivo* or in a cell-free ribosomal system, to yield coelenterazine. One suitable pre-coelenterazine peptide has a modified amino acid sequence of GFP in which $R^{65}$ is Tyr, $R^{80}$ is Gln, $R^{100}$ is Phe, $R^{108}$ is Thr, $R^{141}$ is Leu, $R^{172}$ is Glu, and $R^{219}$ is Val. In another suitable peptide, $R^{80}$ is Gln, $R^{100}$ is Tyr, $R^{108}$ is Ser, and $R^{141}$ is Met, $R^{172}$ is Glu, and $R^{219}$ is Ile. In further suitable pre-coelenterazine peptides, any of $R^{229}$ through $R^{238}$ may be omitted or replaced without detriment to the ability of the pre-coelenterazine to release coelenterazine.

These peptides may be synthesized by a suitable method such as by exclusive solid phase techniques, by partial solid-phase techniques, by fragment condensation by classical solution phase synthesis, or by recombinant DNA techniques.

<u>In a second embodiment</u>, the invention provides polynucleotides, each of which comprises one or more sequences of nucleotide bases collectively encoding a modified amino acid sequence of a GFP of *A. victoria* comprising in which $R^{65}$ is Tyr. Certain of these polynucleotides include at least $R^{1}$ through $R^{228}$. In the polynucleotides of this embodiment, the nucleotides encoding for several amino residues may vary: $R^{80}$ may be Gln or Arg, $R^{100}$ may be Phe or Tyr, $R^{108}$ may be Thr or Ser, $R^{141}$ may be Leu or Met, $R^{172}$ may be Glu or Lys, and $R^{219}$ may be Val or Ile.

The one or more sequences of bases collectively encoding the pre-coelenterazine peptide in these polynucleotides may be uninterrupted by non-coding sequences. For example, the polynucleotide may suitably be a cDNA encoding for a modified GFP gene of *A. victoria* in which the nucleotides for amino acid 65 have been mutated from TCT to TAT. This polynucleotide is, in the present invention, denominated gfp(C197A) to indicate the mutation in the cDNA of GFP at nucleotide 197 from C to A. Further, these polynucleotides may suitably include the incorporation of codons "preferred" for expression by selected mammalian or non-mammalian hosts.

These polynucleotides may comprise further encoding sequences. Thus, one polynucleotide comprises, in addition to the sequence encoding the pre-coelenterazine peptide, one or more sequences of nucleotide bases collectively encoding the amino acid sequence of a luciferase compatible with coelenterazine. Another polynucleotide comprises, in addition to the pre-coelenterazine encoding sequence, one or more sequences of nucleotide bases collectively encoding the amino acid sequence of aequorin.

When in an expression vector, all of the above polynucleotides may further comprise, 5′ or 3′ of the one or more polypeptide encoding sequences, one or more appropriate regulatory elements controlling expression of these sequences. Depending on the type of expression vector and regulatory element used, one regulatory element may be operatively linked to one or more than one encoding sequences. One expression vector comprises a polynucleotide comprising sequences of nucleotide bases collectively encoding a modified *A. victoria* GFP wherein $R^{65}$ is Tyr and one or more sequences of nucleotide bases which encode at least one regulatory element operatively linked to the sequences encoding the pre-coelenterazine peptide. Suitably, the regulatory element is from a gene encoding other than GFP.

Suitably regulatory elements, which are well known to those skilled in the art, include promoters and enhancers. The regulatory elements are operatively linked to a polypeptide encoding sequence when they control, *i.e.*, enable, modulate, activate or deactivate, directly or indirectly, the expression of these sequences.

When in an expression vector, the polynucleotide carrying one or more appropriate regulatory elements may further carry one or more further sequences of bases which collectively confer resistance to an antibiotic, when the polynucleotide is expressed in an organism. One suitable expression vector comprises gfp(C197A); another is plasmid TU#132.

In one expression vector, the regulatory element is a promoter selected from the group consisting of promoters from a P450 gene, a promoter activated by a heavy metal, and a promoter from a gene encoding a stress protein.

Another expression vector comprises, in addition to the sequence encoding the pre-coelenterazine peptide, one or more sequences of nucleotide bases collectively encoding a luciferase compatible with coelenterazine. This expression vector may have one or more sequences of nucleotide bases encoding a further regulatory element operatively linked to the sequences of nucleotide bases encoding said luciferase. If desired, the regulatory element operatively linked to the sequences encoding pre-coelenterazine peptide may be the same as, or different from, the further regulatory element operatively linked to said one or more sequences encoding luciferase.

An expression vector comprising a sequence encoding pre-coelenterazine peptide may also further comprise one or more sequences of nucleotide bases collectively encoding aequorin.

There are further embraced in this second embodiment, a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding the amino acid sequence of $R^1$ through $R^{69}$ of said pre-coelenterazine peptide, as well as an oligonucleotide encoding the amino acid sequence of amino acid residues 64 through 69 of a pre-coelenterazine peptide. One such oligonucleotide has the nucleotide sequence TTC TAT TAT GGT GTT CAA.

<u>In a third embodiment</u> of the present invention, a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding the modified amino acid sequence of a GFP of *A. victoria* in which $R^{65}$ is Tyr, including those in which $R^{80}$ is Gln or Arg, $R^{100}$ is Phe or Tyr, $R^{108}$ is Thr or Ser, $R^{141}$ is Leu or Met, $R^{172}$ is Glu or Lys, and $R^{219}$ is Val or Ile is introduced as exogenous polynucleotide material into an organism. In certain of these polynucleotides, the sequence of nucleotide bases collectively encode at least residues $R^1$ through $R^{228}$ of the modified GFP.

Transformation of these cells may be performed by techniques well known to persons having skill in the art with appropriate expression vectors, e.g., plasmid TU#132. Methods to introduce exogenous genetic material into a cell are well-known in the art. For example, exogenous DNA material may be introduced into the cell by calcium phosphate precipitation technology. Other technologies, such as the retroviral vector technology, electroporation, lipofectiom and other viral vector systems such as adeno-associated virus system, or microinjection may be used. For example, a bacteriophage carrying a polynucleotide encoding the pre-coelenterazine peptide may be used to infect a particular type of bacteria. The infection may be subsequently detected by lysing said bacteria or its progeny in a medium containing a compatible luciferase. Accordingly, by using bacteriophages modified to carry such a polynucleotide, the presence in a sample of particular types of bacteria may be detected. Similarly, a eucarycotic virus carrying the polynucleotide encoding the pre-coelenterazine peptide

may infect a specific cell type. This infection may also be easily detected by lysing said cells in a medium containing a compatible luciferase.

Organisms into which these polynucleotides may be introduced include bacterial cells, yeast cells, fungal cells, insect cells, nematode cells, plant or animal cell. Suitable bacterial cells include *E. coli* BL21 (DE3)Lys S and *E. coli* BLR (DE3).

All of these organisms may additionally be transformed with a second polynucleotide or expression vector comprising one or more sequences of nucleotide bases collectively encoding a luciferase compatible with coelenterazine or collectively encoding aequorin. Squid giant neuron cells transformed with an expression vector encoding pre-coelenterazine peptide are suitable for transformation with an expression vector encoding aequorin.

In a fourth embodiment, there is provided a method of synthesizing a peptide comprising the modified amino acid sequence of a GFP of *A. victoria* including at least $R^1$ through $R^{228}$ in which $R^{65}$ is Tyr, including those in which, $R^{80}$ is Gln or Arg, $R^{100}$ is Phe or Tyr, $R^{108}$ is Thr or Ser, $R^{141}$ is Leu or Met, $R^{172}$ is Glu or Lys, and $R^{219}$ is Val or Ile. This method comprises incubating a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding an amino acid sequence of such a peptide in the presence of means for effecting expression of the polynucleotide under conditions favorable for the expression of the polynucleotide.

The step of incubating the polynucleotide may be preceded by transforming an organism with the polynucleotide, and in which the means for effecting expression of the polynucleotide is the transformed organism.

In one variant of this method, there is provided a method of synthesizing coelenterazine comprising synthesizing a pre-coelenterazine peptide according to the method in Claim 21, and isolating coelenterazine

from said means. This embodiment concerns an efficient method for expression of coelenterazine such that large amount thereof may be produced. Methods to collect or isolate coelenterazine are well-known and therefore, coelenterazine may be isolated easily.

Any of the organisms identified above is suitable; transformation may be by any procedures deemed appropriate by those skilled in the art. The step of incubating the polynucleotide may be performed by culturing the transformed organism for one or more generations under conditions favorable to growth of the transformed organism and to expression of the polynucleotide, and the step of isolating coelenterazine may be performed by lysing the progeny of the cultured transformed organism to form a cell-free extract, and isolating coelenterazine from this extract.

In this method, the means for effecting expression of said polynucleotide may be *E. coli* strain BL21(DE3)Lys S (Studier and Moffatt, J.Mol.Biol. 189 113 (1986), incorporated herein by reference) or *E. coli* BLR (DE3), transformed with an expression vector comprising, 5′ or 3′ of said one or more sequences of nucleotide bases collectively encoding the amino acid sequence of pre-coelenterazine peptide, one or more appropriate regulatory elements which collectively enable expression of said polynucleotide; and one or more sequences of bases which collectively confer resistance to an antibiotic upon an organism. One suitable transformed *E. coli* BL21(DE3)Lys S is *E. coli* SMC2 (ATCC Accession No. 69553).

Alternatively, the means for effecting expression of said polynucleotide when it is a polyribonucleotide may be a cell-free aqueous translation system known to those skilled in the art.

The method may further comprise the step of converting isolated coelenterazine to a stable form, luciferyl sulfate, as for example by incu-

bating isolated coelenterazine with a luciferin sulfokinase. The luciferin sulfokinase may suitably be derived from the organism *R. reniformis*.

This embodiment further includes purified coelenterazine and purified luciferyl sulfate synthesized by one of the above methods.

A fifth embodiment provides assays employing coelenterazine bioluminescence as an indicator. Several of these assays employ a further polynucleotide or expression vector having one or more sequences of nucleotide bases collectively encoding either the amino acid sequence of apo-aequorin or a luciferase compatible with coelenterazine; or an organism transformed with one of these polynucleotides. (A luciferase is compatible with coelenterazine if, when combined in an aqueous medium with coelenterazine, it generates bioluminescence at or about 480 nm.) One such suitable luciferase is that isolated from the sea pansy *R. reniformis* and encoded in the polynucleotide disclosed in Lorenz et al., Proc.Nat.Acad.Sci., 88, 4438-4442, 1991, incorporated herein by reference. A suitable polynucleotide containing sequences which encode the apopeptide of aequorin is available commercially from Sealite Corp. of Atlanta, Georgia.

As with the above polynucleotides, the luciferase polynucleotide may further comprise appropriate regulatory elements and sequences conferring antibiotic resistance, and even the polynucleotide comprising sequences further may comprise one or more sequences of nucleotide base collectively encoding the amino acid sequence of the pre-coelenterazine peptide.

Upon expression of the pre-coelenterazine peptide and luciferase genes, organisms which do not naturally bioluminesce at 480 nm exhibit bioluminescence at or about 480 nm. These assays may be employed for a variety of uses, as for example to detect the expression of certain genes or proteins of interest in cells; detecting increased levels of intracellular calcium ion; or detecting the presence of $O_2$ in an anaerobic system.

The above-described cells and living organisms are useful to detect effects of external stimuli to the regulatory element. The stimuli may have direct or indirect effects on the regulatory element. Such effects will be detectable through either the induction of expression and production of the pre-coelenterazine peptide which, upon exposure to a compatible luciferase, results in bioluminescence; or through the switching off the expression of the pre-coelenterazine peptide.

These cells and organisms may be used to detect the presence of certain molecules in various kinds of biological samples such as blood, urine or saliva. By operatively linking a regulatory element which is affected by the molecule of interest to a polynucleotide sequence encoding the pre-coelenterazine peptide, the presence of the molecules will affect the regulatory element which in turn will affect expression of the pre-coelenterazine peptide. Detection of these molecules may be used for diagnostic purposes. An example of such a molecule is a hormone.

These assays may further be used to localize a protein of interest in a cell, both described in Chalfie et al., *supra*. More particularly, this embodiment provides a method for selecting cells expressing a protein of interest, or for detecting expression of a gene of interest. Thus, the method for selecting cells expressing a protein of interest comprises introducing into cells the polynucleotide encoding the pre-coelenterazine peptide, and a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding said protein of interest. These cells are then cultured under conditions permitting expression of the pre-coelenterazine peptide and the protein of interest. The cells are then examined for expression of coelenterazine; those which express coelenterazine are thereby selected cells expressing the protein of interest.

There are several conventional means by which one may identify cells expressing coelenterazine. One may plate out the cultured cells and grow

colonies from each. Samples taken from each colony may be cultured in turn, and lysed in a medium containing a luciferase compatible with coelenterazine. The exhibition of bioluminescence in this medium confirms the expression of the pre-coelenterazine peptide, and thus of the protein of interest. Alternatively, one could also, at the time the cells are transformed with the polynucleotide encoding pre-coelenterazine peptide, further transform the cells with a polynucleotide comprising one or more sequences which collectively encode for a compatible luciferase. Cells exhibiting bioluminescence could then be isolated by means of well known to persons skilled in the art.

The above cells and organisms are also useful in methods for detecting expression of a gene of interest. This method comprises introducing into a cell a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding a regulatory element operatively linked to the gene of interest, as well as a polynucleotide which encodes for the pre-coelenterazine peptide, such that the regulatory element of the gene controls expression of pre-coelenterazine peptide, The cells are then cultured in conditions permitting expression of the gene of interest and of the pre-coelenterazine peptide. One then detects the expression of coelenterazine in the cell by means well known to the art, thereby indicating the expression of the gene in the cell.

A method for detecting increased levels of intracellular calcium ion comprises the steps of culturing an organism transformed with a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding a pre-coelenterazine peptide, and a second polynucleotide comprising one or more sequences of nucleotide bases collectively encoding aequorin. The organism is cultured under conditions favorable to its growth and to expression of the pre-coelenterazine and aequorin peptides and monitored for exhibition of bioluminescence. The exhibition of biolumin-

escence by the cultured organisms indicates that intracellular levels of calcium ion have risen above normal cytoplasmic levels.

Aequorin and coelenterazine complex into a heterotetramer in coelenterates of the order hydrozoa; they also do so when synthesized in the transformed organism. The usual intracellular levels of calcium ion of approximately $10^{-7}$ M are insufficient to trigger the heterotetramer form to exhibit bioluminescence. However, intracellular concentrations of calcium ion rise following opening of ion channels in the plasma membrane, damage to the cell membrane, or release of calcium ion into the cytoplasm from its intracellular repositories, the mitochondria or endoplasmic reticulum. Following these events, intracellular calcium ion concentrations may rise to levels of $10^{-5}$-$10^{-3}$ M. These levels are more than adequate to trigger the heterotetramer to bioluminesce. Accordingly, this method permits one to monitor the frequency of calcium ion concentration increases and to evaluate physiological events which accompany these increases. Suitably the cell which is transformed is the squid giant neuron.

This assay improves on the conventional assays employing intracellular aequorin (described in Grynkiewicz G. et al., J.Biol.Chem. 260 3440-50 (1985); Tsien R.Y. et al., Trends Biochem.Sci. 11 450 (1986); and Gilkey J. C. et al., J.Cell.Biol. 76 448-466 (1978), all incorporated herein by reference) because the level of aequorin and coelenterazine in the cell may be more accurately controlled by appropriate selection of the regulatory elements.

A method for detecting the presence of $O_2$ leaks into an anaerobic system comprises the steps of culturing an organism transformed with a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding a pre-coelenterazine peptide, and a second polynucleotide comprising one or more sequences of nucleotide bases collectively encoding a compatible luciferase. The organism is cultured under

conditions favorable to growth of said transformed organism and favorable to expression of pre-coelenterazine and the luciferase, where the organism is a facultative anaerobe; and monitoring the culture for exhibition of bioluminescence. Any exhibition of bioluminescence by the culture indicates that $O_2$ has leaked into the anaerobic system.

In a sixth embodiment, the invention provides an organism transformed with two polynucleotides. The first polynucleotide comprises one or more sequences of nucleotide bases collectively encoding a pre-coelenterazine peptide comprising a modified *A. victoria* GFP having an amino acid sequence in which $R^{65}$ is Tyr. The second polynucleotide comprises one or more sequences of nucleotide bases collectively encoding an amino acid sequence for a luciferase peptide compatible with coelenterazine. One of the polynucleotides has a mutation which precludes a bioluminescent interaction between their expression products. The mutation is desirably reversible upon exposure of the transformed organism to a mutagen; reversal of the mutation enables a bioluminescent interaction between the two expression products.

Any of the organisms identified above is suitable; transformation may be by any procedure deemed appropriate by those skilled in the art. The transformed organism may be employed in an assay to detect mutagenesis, as in a modified "Ames test." Ames et al., Proc.Nat.Acad.Sci., 70, 782-786 and 2281-2285 (1973), incorporated herein by reference.

This embodiment further provides a method of detecting mutagenesis caused by a chemical compound suspected of being a mutagen. The method comprises the steps of transforming a population of organisms with both of the first and second polynucleotides described above; growing a culture of said transformed organisms through one or more generations in a nutrient medium comprising said chemical compound; and measuring the

bioluminescence of said culture and comparing said bioluminescence to that from a culture of non-transformed mutagenized control organisms.

## DESCRIPTION OF THE DRAWING

Figure 1 is the print out of a luminometry test of coelenterazine isolated from *E. coli* SMC2.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Abbreviations and Conventions. The usual conventions for indicating polypeptides (written with N-terminal to left, C-terminal to right) and polynucleotides (written 5' to left, 3' to right) are followed herein. The residues of the pre-coelenterazine peptide are numbered according to Prasher, et al., Gene 111, 229-233, 1992, (incorporated herein by reference) beginning with $R^1$ at the N-terminal and proceeding sequentially toward the C-terminal.

Any DNA disclosed as an individual single-stranded DNA also is considered to disclose the double-stranded DNA forming the same, as well as RNA equivalent thereto.

Applicants state at several points herein that methods, techniques, organisms, and various means for carrying out identified procedures, all of which are well known in the art, may suitably be used. This statement is not to be interpreted that every possible alternative means is equally desirable and effective; the choice among these alternative techniques, organisms and combinations thereof is left to the skill and discretion of one skilled in the art. Conversely, Applicants' statements herein that certain specific techniques and organisms may suitably be used is not to be interpreted that these specified techniques or organisms, or certain combinations thereof, are particularly preferred. The identification of these techniques and organisms is merely exemplary.

A. PRE-COELENTERAZINE PEPTIDE

The primary structure of the pre-coelenterazine peptide is substantially based on that of the GFP of *A. victoria* appearing in Prasher et al., *supra*, except that $Ser^{65}$ appearing in the chromogenic sequence of amino acyl residues of apo-GFP, i.e., $Phe^{64}$-$Ser^{65}$-$Tyr^{66}$-$Gly^{67}$-$Val^{68}$-$Gln^{69}$, is replaced by Tyr.

Variations are seen in the amino acid sequence of wild *A. victoria* jellyfish GFP at several residues: $R^{80}$ may be Gln or Arg, $R^{100}$ may be Phe or Tyr, $R^{108}$ may be Thr or Ser, $R^{141}$ may be Leu or Met, $R^{172}$ may be Glu or Lys, and $R^{219}$ may be Val or Ile. These same replacements may be made in the pre-coelenterazine without substantial prejudice. The length of the pre-coelenterazine peptide may also be subject to minor differences in length: i.e., the primary sequence of the peptide may slightly exceed or fall short of 238 amino acyl residues.

This length variation may arise when the peptide is derived from a polypeptide having 5′ or 3′ termini to which "sticky end" nucleotide sequences have been added by procedures known to persons skilled in the art to facilitate insertion of the polynucleotide into an appropriate vector.

For purposes of this disclosure, additional amino acid residues in the pre-coelenterazine peptide are indicated, not by altering the numbering of residues, but by considering the extra residues as one with either $R^1$ or $R^{238}$. Thus, in the peptide of SEQ ID NO: 1 $R^1$ is Xaa is methionyl-alanine.

Conversely, when one or more amino acid residues are deleted due to the introduction of restriction sites in the cDNA, the amino acid residues are numbered according to the number they would have held in GFP. Thus, if $R^1$ through $R^3$ were omitted, the N-terminal Gly would nevertheless be numbered $R^4$ in the truncated peptide.

All of these pre-coelenterazine peptides may be synthesized by a suitable method such as by exclusive solid phase techniques, by partial solid-phase techniques, by fragment condensation or by classical solution phase synthesis. For example, the techniques of exclusive solid-phase synthesis are set forth in the textbook "Solid Phase Peptide Synthesis", J.M. Stewart and J.D. Young, Pierce Chem. Company, Rockford, 111, 1984 (2nd. ed.), and M. Bodanszky, "Principles of Peptide Synthesis", SpringerVerlag, 1984. The peptides may suitably be prepared using solid phase synthesis, such as that generally described by Merrifield, J.Am.Chem.Soc., 85, p. 2149 (1963), although other equivalent chemical syntheses known in the art may also be used as previously mentioned.

Alternatively, each of these peptides may be made using recombinant DNA techniques. This may be done, for example by generating a polynucleotide which encodes, according to the genetic code of chromosomal DNA, the amino acid sequence of the desired pre-coelenterazine peptide. This polynucleotide, introduced into an expression vector, may be expressed *in vitro* or *in vivo*. The polynucleotide may be generated by procedures well known in the art, e.g., by DNA or RNA synthesis techniques and/or devices or by introducing one or more point mutations into the gene for GFP. Other suitable methods for this and other recombinant DNA techniques are discussed in Sambrook, Molecular Cloning: A laboratory manual, 2nd Ed., Cold Spring Harbor Laboratory Press (1989).

The invention further concerns a peptide derived from the pre-coelenterazine peptide in which one or both $R^{65}$ and $R^{66}$ are dehydroTyr. This peptide may be generated by synthesizing the pre-coelenterazine peptide *in vivo*.

B. POLYNUCLEOTIDE ENCODING THE PRE-COELENTERAZINE PEPTIDE

In a second embodiment, the invention provides a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding

the modified amino acid sequence of a GFP of *A. victoria* in which $R^{65}$ is Tyr.

The polynucleotide may also be comprised of one or more sequences of nucleotide bases collectively encoding the modified amino acid sequence of a GFP of *A. victoria* in which $R^{65}$ is Tyr, and further, in which $R^{80}$ may be Gln or Arg, $R^{100}$ may be Phe or Tyr, $R^{108}$ may be Thr or Ser, $R^{141}$ may be Leu or Met, $R^{172}$ may be Glu or Lys, and $R^{219}$ may be Val or Ile. Two suitable polynucleotides comprise one or more sequences of nucleotide bases collectively encoding for such a pre-coelenterazine in which $R^{80}$ is Gln, $R^{100}$ is Phe, $R^{108}$ is Thr, $R^{141}$ is Leu, $R^{172}$ is Glu, and $R^{219}$ is Val; or in which $R^{80}$ is Gln, $R^{100}$ is Tyr, $R^{108}$ is Ser, $R^{141}$ is Met, $R^{172}$ is Glu and $R^{219}$ is Ile.

These polynucleotides may be composed of either DNA or RNA, and may be either single or double stranded.

These polynucleotides may have, in the one or more sequences collectively encoding the pre-coelenterazine peptide, any nucleotide sequence which encodes one of the pre-coelenterazine peptides under the chromosomal genetic code. This code is degenerate; thus, the Arg at residue 109 may be encoded by nucleotide bases as CGT, but could, under the code, equally be CGC, CGA, CGG, AGA or AGG, and still encode for Arg. Similarly, nucleotide bases encoding Tyr as TAT could equally be TAC and still encode Tyr for residue 65 in the peptide. Alternatively, the polynucleotide may be a cDNA (or RNA equivalent) which includes, in addition to the nucleotides for $Ser^{65}$ being altered from TCT to TAT, mutations encoding for $R^{80}$ may be Gln or Arg, $R^{100}$ may be Phe or Tyr, $R^{108}$ may be Thr or Ser, $R^{141}$ may be Leu or Met, $R^{172}$ may be Glu or Lys, and $R^{219}$ may be Val or Ile. Thus, for example, the polynucleotide may include one or more of the following base mutations: the bases encoding $Gln^{80}$ may be altered from CAG (for Gln) to CGG (for Arg); the bases encoding $Phe^{100}$ may be altered from TTC to TAC (Tyr); bases for $Thr^{108}$ may be altered from

ACA to AGA (for Ser); bases for $Leu^{141}$ may be altered from CTG to ATG (for Met); bases for $Glu^{172}$ may be altered from GAA to AAA (for Lys). All these polynucleotides which encode for the pre-coelenterazine peptides are found, upon expression, to direct the synthesis of one of the pre-coelenterazine peptides, and hence are embraced in the invention.

To generate a polynucleotide having these encoding sequences, one may introduce one or more point mutations into the cDNA for the GFP of *A. victoria* in the cDNA described in Prasher, et al., *supra*, using *in vitro* mutagenesis methods well known to those skilled in the art. If desired, one may make further mutations to effect changes at the nucleotides encoding amino acid residues 80, 100, 108, 141 and 219 as desired.

The abbreviation gfp(C197A) is used herein to designate the cDNA sequence of SEQ ID NO: 2. Nucleotides 1 through 717 of gfp(C197A) encode the amino acid sequence of the pre-coelenterazine peptide, *i.e.*, the modified *A. victoria* GFP of SEQ ID NO: 1. The alteration of $R^{65}$ from Ser to Tyr is effected by the mutation of nucleotide 197 from C to A. The polydeoxyribonucleotide sequence of gfp(C197A) appears in SEQ ID NO:2, where bases 1 through 717 are encoding bases; the mutated base 197 is A; and the triplet codon formed by bases 718 through 720 form the stop codon.

All of the above polynucleotides may further comprise, 5′ or 3′ of said one or more sequences of bases, one or more appropriate regulatory elements which collectively enable expression of said one or more sequences of bases encoding said pre-coelenterazine peptide.

Suitably, the regulatory element may be a promoter. Suitable promoter elements include a promoter activated by heavy metal (e.g. the one described in Freedman, et al. J. Biological Chemistry, 268: 2554, 1993, incorporated herein by reference); a P450 promoter (e.g. the cytochrome

P450); or a promoter for a stress protein, (e.g., described in Stringham, et. al., Molecular Biology of the Cell, 3: 221, 1992), one of said stress proteins being a heat-shock protein. Other suitable promoters include that of the arabinose operon (phi 80 dara) or the colicin E1, galactose, alkaline phosphatase or tryptophan operons. Similarly the ADH system may be employed to provide expression in yeast. Alternatively, the regulatory element may be an enhancer.

The regulatory elements are operatively linked with the polypeptide comprising one or more sequences of nucleotide bases collectively encoding an amino acid sequence of a pre-coelenterazine peptide; i.e., the regulatory elements are placed on the polynucleotide 5' or 3' of the one or more sequences suitable to enable expression of the sequences.

Polynucleotides which bear one or more of such regulatory elements may be used in transforming organisms, as when suitably the polynucleotide is included in an expression vector. The regulatory elements are selected for compatibility with the organism into which the polynucleotide is to be incorporated by transformation, i.e., the regulatory elements are those which may be recognized by the transformed organism or cell and which will aid in controlling the expression of said polynucleotide in the transformed organism.

Thus when the organism to be transformed is *E. coli*, the regulatory element may be a promoter (e.g., the T7, the SP6 or *lac* promoter); or transcription initiation sequences for ribosome binding (e.g. the Shine-Delgarno sequence and the start codon AUG). When the organism to be transformed is eucaryotic, the regulatory elements may include a heterologous or homologous promoter for RNA polymerase II and/or a start codon AUG. For example, when the target of transformation is a mammalian cell, the regulatory element may be a promoter (e.g. the SP 40 or the bovine papilloma virus promoter). Suitable regulatory elements for use in other

microbe, or in animal, or plant cells may be selected according to criteria well known to persons having skill in the art. All of these regulatory elements may be obtained commercially (individually or incorporated into a vector) or assembled by methods well known in the art.

Where the polynucleotide carries one or more appropriate regulatory elements, there may further be present one or more further sequences of bases which collectively confer resistance to an antibiotic when the polynucleotide is expressed in an organism. Such genes for antibiotic resistance are desirable components for expression vectors since they facilitate identification of transformed cells grown in the presence of antibiotics, and exert a continual pressure on the transformed organisms to retain and express the expression vectors. One polynucleotide suitable for use in transforming bacteria *E. coli* is plasmid TU#132.

All of the above polynucleotides may be synthesized by known methods. The polynucleotide may be generated by procedures well known in the art, e.g., by DNA or RNA synthesis techniques and/or devices or by introducing one or more point mutations into the gene for GFP. Other suitable methods for this and other recombinant DNA techniques are discussed in Sambrook, *supra*.

Thus one may use a DNA synthesizing device to construct the entire polynucleotide or to synthesize several fragments of a polynucleotide and ligate these together. This is a laborious process for polynucleotides, and thus it is usually preferable to generate polynucleotides by other means known to those skilled in the art.

The embodiment further includes a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding the amino acid sequence of $R^1$ through $R^{69}$ of the pre-coelenterazine peptide, as well as an oligonucleotide comprising nucleotide bases encoding the sequence of

amino acid residues 64 through 69 of the pre-coelenterazine peptide. One such oligonucleotide has the nucleotide sequence TTC TAT TAT GGT GTT CAA. Such poly- and oligonucleotides are useful as probes for homologous sequences of DNA or RNA. The attachment of a chemical label to such probes facilitates locating the probes in a test system. Suitable labels include radioisotopes (e.g. $^{32}P$, $^{35}S$, $^{125}I$), fluorescent compounds, or other well known labels (e.g. biotin) covalently linked to the poly- or oligonucleotide.

Although probes are normally used with a detectable label that allows easy identification, these poly- and oligonucleotides are also useful in unlabeled form, both as precursors of labeled probes and for use in methods that provide for direct detection of double-stranded DNA or DNA/RNA.

C. ORGANISMS TRANSFORMED WITH POLYNUCLEOTIDE ENCODING PRE-COELENTRAZINE PEPTIDE

A third embodiment of the present invention is an organism transformed with a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding the modified amino acid sequence of a GFP of *A. victoria* in which $R^{65}$ is Tyr, $R^{80}$ is Gln or Arg, $R^{100}$ is Phe or Tyr, $R^{108}$ is Thr or Ser, $R^{141}$ is Leu or Met, $R^{172}$ is Glu or Lys, and $R^{219}$ is Val or Ile. Organisms which may suitably be transformed with such polynucleotides include bacterial cells, yeast cells, fungal cells, insect cells, nematode cells, plant or animal cell. Suitable animal cells include, but are not limited to Vero cells, HeLa cells, Cos cells, CV1 cells and various primary mammalian cells.

Transformation of these cells may be performed by techniques well known to persons having skill in the art. Thus, for instance, transformation of yeast and plant cells must be preceded by treatment of the cells to remove the rigid cell wall, as by treatment with a digestive enzyme; the

resulting spheroplasts and chloroplasts readily take up polynucleotide plasmids and upon return to growth medium regenerate their cell walls.

One suitable transformed organism is *E. coli* SMC2, an *E. coli* of the Strain BL21 (DE3)Lys S which has been transformed with Plasmid TU#132. *E. coli* SMC2 and plasmid TU#132 were deposited on February 4, 1994 with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD 20852, USA, under the provision of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Title 37 Section 1.801 et seq. of the Code of Federal Regulations, and accorded ATCC Accession Nos. 69553 and 75666 respectively.

It is noted that the deposited material is not considered to be essential to the practice of the claimed invention and that the grant of admission to the depository to distribute samples of the biological material does not constitute an express or implied license to practice the invention claimed in any patent issuing from the instant application or from any continuation, divisional or reissue application thereof.

Another suitable cell is *E. coli* strain BLR (DE3) (A. Roca, University of Wisconsin, cited in the Novogen Catalogue). When this strain is transformed with an expression vector comprising a polynucleotide having one or more sequences of nucleotide bases collectively encoding a pre-coelenterazine peptide, the resulting strain produces coelenterazine more stably. The stability of this production is believed to be due to the reduced recombination of the host. *E. coli* strain BLR (DE3) may suitably be transformed with an expression vector based on pET11.

D. METHOD OF SYNTHESIZING PRE-COELENTERAZINE AND COELENTERAZINE

The fourth embodiment of the present invention comprises a method of synthesizing a pre-coelenterazine peptide comprising a modified *A. victoria* GFP having an amino acid sequence in which $R^{65}$ is Tyr. The method comprises incubating a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding an amino acid sequence of such a peptide in the presence of means for effecting expression of the polynucleotide under conditions favorable for the expression of the polynucleotide. These means may be an *in vitro* transcription/translation system or an organism transformed with the polynucleotide.

In this method, the means for effecting expression of the polynucleotide may be either an *in vitro* cell-free translation system or an organism which has been transformed with the polynucleotide, being viable and in a medium containing assimilable sources of carbon, nitrogen, and inorganic substances. One suitable means for effecting expression *in vivo* is *E. coli* SMC2.

When the means for effecting expression is such a transformed organism, the polynucleotide may comprise one or more appropriate regulatory elements, and one or more sequences of bases which collectively confer resistance to an antibiotic upon the transformed organism. One suitable polynucleotide is Plasmid TU#132.

There is further provided a method of synthesizing coelenterazine comprising synthesizing the pre-coelenterazine peptide according to the above method and isolating coelenterazine from the means for effecting expression of the polynucleotide. This embodiment provides an efficient method for expression of coelenterazine such that large amounts of the compound may be produced. Methods to isolate expressed protein have been well-known and therefore, coelenterazine may be isolated easily.

In this method, the means for effecting expression may be an organism capable of expressing the polynucleotide, i.e., one transformed or transformable with the polynucleotide when subjected to conditions favorable to transformation in the presence of the nucleotide. The transformed organism and the transformable organism (once the latter is transformed) is cultured for one or more generations under conditions favorable to growth of said organism and to expression of the polynucleotide. In this method, the step of isolating coelenterazine is performed by lysing the progeny of said cultured transformed cells to form a cell-free extract, and isolating coelenterazine from said extract.

To maximize expression of the pre-coelenterazine peptide, the sequence flanking the translation initiation codon may be modified (reviewed by Kozak, 1984), compilation and analysis of sequences upstream from the translation start site in eucaryotic mRNA's Nucl. Acids Res. 12:857-872, incorporated herein by reference. A sequence may then be generated to produce higher amounts of the pre-coelenterazine peptide. In addition, artificial introns may be introduced so as to increase the production of the protein.

The transformed cell selected to express the polynucleotide encoding the pre-coelenterazine peptide also affects the level of expression. Expression may also be boosted by employing said method, as a means for effecting expression of the polynucleotide, a cell selected from the group consisting of *E. coli* SMC2 (ATCC Accession No. 69553) and *E. coli* BLR (DE3) transformed with the pET3a expression vector described above. These cells are suitably cultured as described in Example III below, optionally in the presence NADP.

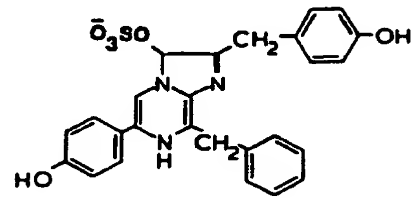
The conditions of growth also may be modified in order to raise the level of pre-coelenterazine produced. When the transformed cells are cultured at 30°C and in the absence of IPTG until they reach log growth

phase. At this stage, when a large number of cells are present, IPTG (suitably 0.5mM to 1mM) is added to induce expression of the polynucleotide encoding the pre-coelenterazine peptide. After inducing for a suitable time, the cells are harvested, and coelenterazine may be collected.

Coelenterazine is a highly labile substance in aqueous based media with a half-life of one to two hours. When it is suspended in methanolic HCl, however it is stable, even at room temperature. Accordingly, the step of isolating coelenterazine from said extract may be performed by adding methanolic HCl to the cell lysate, mixing and removing the suspended cell solids.

It is well known that coelenterazine may be stabilized in aqueous media by being modified from its keto form (Formula III) to its enol form in luciferyl sulfate. The enol sulfate form of coelenterazine has conventionally been termed luciferyl sulfate and this term is employed herein. Luciferyl sulfate has the structure shown in Formula V:

Formula V



This modification to luciferyl sulfate is carried out by incubating coelenterazine with the enzyme luciferin sulfokinase and 3',5'-diphosphoadenosine. The conversion of luciferyl sulfate to luciferin is 3,5-diphosphoadenosine-linked. This may be accomplished by incubating the isolated coelenterazine with a compatible luciferin sulfokinase, suitably a luciferyl sulfokinase derived from a coelenterate such as *R. reniformis*. Methods well known in the art may be used to isolate and purify luciferyl sulfokinase, e.g., methods described in Cormier et al., J.Cell.Physiol. 81, No. 2, 291-297 (1973), incorporated herein by reference; Hori et al., Biochim.

Biophys.Acta 256: 638-644 (1972), incorporated herein by reference; and Karkhanis and Cormier, Bioch. 10:317-326 (1971), incorporated herein by reference. Luciferin sulfokinase may be isolated from this organism according to the procedures described in Cormier et al., Bioch. 9 1184-1189, 1970, incorporated herein by reference.

Purification of the coelenterazine from the lysate may be accomplished using chromatography procedures well known in the art. Such procedures include size exclusion chromatography, column chromatography and high performance HPLC using one or more reverse phase HPLC procedures. Combinations of such chromatography methods may also be employed. In this manner coelenterazine is obtained in purified form.

Coelenterazine synthesized by these methods may be characterized by one or more of the following methods: HPLC, emission spectroscopy, and mass spectroscopy. Performance of these tests upon coelenterazine synthesized according to the above methods demonstrates that the coelenterazine has the same chromatographic profile as natural coelenterazine; that the synthetic coelenterazine emits blue light at 480 nm; and that it has a mass spectroscopy profile nearly identical to that of natural coelenterazine.

Accordingly, in a further embodiment of the invention, there is provided purified coelenterazine and luciferyl sulfate made by the above methods. The production process of these compounds is carried out in a conventional manner. Transformed bacterial cells are resuspended in a suitable known buffer solution, followed by lysing the bacterial cells in a conventional manner such as ultrasonic wave treatment and/or enzyme treatment, and obtaining the supernatant by means of centrifugation.

E. IMPROVED ASSAY FOR DETECTING MUTAGENESIS

In another embodiment, the invention provides an organism transformed with two polynucleotides, one comprising one or more sequences of nucleotide bases collectively encoding an amino acid sequence for a luciferase peptide compatible with coelenterazine. The second polynucleotide comprises one or more sequences of bases collectively encoding the amino acid sequence of said pre-coelenterazine peptide.

One of these polynucleotides has a mutation which prevents its expression product from having a bioluminescent interaction with any coelenterazine. This mutation, which may suitably be an insertion, a duplication, a translocation, mis-sense, or a reading shift mutation, is desirably reversible upon exposure to a mutagen. Thus, should exposure to a mutagen reverse the mutation in the second polynucleotide, the organism will express the first and second polynucleotides, thus generating coelenterazine and an active luciferase compatible therewith.

Since reversal of the mutation results in bioluminescence, the organism may be employed in a mutagenesis assay. Therefore, the embodiment further provides a method of testing the mutagenicity of a chemical compound, comprising: a) transforming a population of organisms with said first and said second polynucleotides; b) growing a culture of said transformed organisms through one or more generations in a nutrient medium comprising said chemical compound; and c) measuring the bioluminescence of said culture and comparing said bioluminescence to that from a culture of non-transformed mutagenized control organisms.

In this method, the rate of mutagenesis may be measured instrumentally, as by subjecting said mutagenized cultures to on-line luminometry, avoiding the Ames test steps of preparation of agar plates and tedium of scoring bacterial colonies on the plates.

In order to halt the growth of non-revertant organisms while performing this method, and to prevent "crowding out" of revertant growth by non-revertants, the expression of coelenterazine and luciferase may be linked to mutant survival. This may suitably be done by placing the mutation on a regulatory element controlling the expression of the transformed organisms' antibiotic resistance or the synthesis of an obligatory co-factor. Suitable locations for such a mutation include the promotor region or the region of repressor binding.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications may be made thereto without departing from the spirit or scope of the invention as set forth herein.

Finally, this invention provides a method for producing fluorescent molecular weight markers comprising: a) linking a DNA molecule encoding a green fluorescent protein with a DNA molecule encoding a known amino acid sequence in the same reading frame; b) introducing the linked DNA molecule of step a) in an expression system permitting the expression of a fluorescent protein encoded by the linked DNA molecule; and c) determining the molecular weight of the expressed fluorescent protein of step b), thereby producing a fluorescent molecular weight marker.

Various expression systems are known in the art. The *E. coli* expression system, one of the commonly used systems is described in the following section.

The determination of molecular weight may be done by comparing the expressed fluorescent protein of step b) with known molecular weight markers. Alternatively, the molecular weight can be predicted by calculation since the linked DNA sequence is known (and so is the amino acid sequence being encoded). In an embodiment, the expressed fluorescent protein is

purified. The purified fluorescent protein can be conveniently used as molecular weight markers.

This invention will be better understood from the Experimental Details which follow. However, one skilled in the art will readily appreciate that the specific methods and results discussed are merely illustrative of the invention as described more fully in the claims which follow thereafter.

<u>HPLC</u>. A Bio-Rad ODS-5S reverse-phase HPLC column (4x250 mm) may be used. High-performance liquid chromatography may be carried out on Waters 600 multisolvent system with Waters 490 programmable wavelength detector and Waters 740 data module. Column eluent is monitored at three wavelengths simultaneously so as to detect coelenterazine and UV absorption contamination in real time. All HPLC runs are performed in acidic methanolic buffer solutions, so the coelenterazine is always in the 370 nm-absorbing form.

<u>Photometric Determination.</u> Bioluminescence may be measured and peak light intensities determined, with a luminometer. Bioluminescence intensity is converted to quanta per second by calibrating the instrument relative to a radioactive $^{14}C$ light standard that emits maximally in the 410 nm region. Routine assays for coelenterazine are performed by rapidly injecting 10$\mu$l of clarified *E. coli* SMC2 cell extract in methanolic-HCl into a vial containing *Renilla* luciferase in 1 ml of luciferase buffer.

<u>Analytical Spectra.</u> A Cary 17-D recording spectrophotometer and a Bausch and Lomb Spectronic 2000 are used interchangeably for fixed-wavelength absorbance measurements or for spectral scans.

<u>Mass Spectroscopy.</u> Coelenterazine isolated from recombinant bacteria may be analyzed by electrospray ionization mass spectrometry and liquid secondary ion mass spectrometry followed by mass spectrometry/

mass spectrometry. This will provide molecular weight and structural information on the isolated coelenterazine. These two tools also provide information on the purity of the cyclic tripeptide as well as of other peptides, if present.

## EXAMPLE I

## *In vitro* MUTAGENESIS

TU#58 (described in US Patent Application Serial No 08/119,678) is treated with NcoI and EcoRI to generate a fragment of the GFP gene. This fragment is replicated by PCR with an oligomeric primer to insert the C197A mutation. The fragment is also treated with primers (NcoI at 5', T3 at 3') to incorporate restriction sites. The primer which incorporates the C197A mutation has the sequence:

CCT GTT CCA TGG CCA ACA CTT GTC ACT ACT TTC TAT TAT G

The "A" base located five bases upstream of the 3' terminal G base end constitutes the C197A mutation.

The replicated fragment containing these mutations is then hybridized and ligated to TU#58 which has been treated with NcoI and EcoRI to produce plasmid TU#132. Sequencing of plasmid TU#132 confirms the incorporation of both endonuclease sites and of the C197A mutation.

## EXAMPLE II

## *In vitro* MUTAGENESIS

An alternative method of generating a point mutation to that set out in Example I is as follows. A synthetic oligonucleotide is used to introduce the mutation. The synthetic oligonucleotide may be synthesized employing a commercially available automatic DNA synthesis apparatus, the product being subject to end-phosphorylation in a conventional manner to obtain a primer.

The oligomer is hybridized to the single strand plasmid TU#132 disclosed in US Patent Application Serial No. 08/119,678 under conditions of low stringency, and subjected to three-stage treatment: treatment at 100° C. for 5 minutes, followed by allowing the resulting material to stand at 30° for 3 minutes and further at 4° C. for 30 minutes to carry out annealing and reacting dXTP (X = G, A, T, C) with Klenow fragment (*E. coli* polymerase) in the presence of T4-ligase to prepare a duplex chain.

## EXAMPLE III
## TRANSFORMATION

*E. coli* of strain BL21(DE3)Lys S is transformed with plasmid TU#132. Transformants are cultured at room temperature and selected on media containing ampicillin (100$\mu$g/ml) and IPTG (0.8mM). Plasmid DNA is isolated from the transformants and analyzed by automated sequence analysis, which confirms the presence of the C197A mutation.

## EXAMPLE IV
## MEASUREMENT OF COELENTERAZINE SYNTHESIZED IN *E. coli* SMC2

The bioluminescence of coelenterazine synthesized by *E. coli* SMC2 cultivated as in Example III is measured on a custom built luminometer. The circuitry of this luminometer is modeled after that described in Blinks et al., Methods in Enzymol., 57 292-328, 1978, incorporated herein by reference; the reaction chamber and shutter assembly of the luminometer are modeled after that described in Levine and Ward, Comp.Bioch.Physiol. 72B, 77-85 1982, incorporated herein by reference.

The column is equilibrated with a starting buffer of 0.1% trifloracetic acid at a flow rate of 1mL/min. Coelenterazine-containing samples are injected onto the column in starting buffer. Five minutes after sample injection, a linear methanol gradient (+1% methanol/min, flow rate 1mL/min) is initiated until all components in the mixture are eluted.

Following each HPLC run, the column is rinsed with 100% methanol. Under these conditions, coelenterazine elutes at 90% methanol.

Peak light intensities of the coelenterazine are determined on this luminometer and bioluminescence intensity is converted to quanta per second by calibrating the instrument relative to a radioactive $^{14}C$ light standard that emits maximally in the 410 nm region.

A paste of *E. coli* SMC2 is lysed by sonication in absolute methanol acidified to 1N with HCl. 10$\mu$l of clarified *E. coli* SMC2 cell extract in methanolic-HCl is rapidly injected into a vial containing *Renilla* luciferase in 1 ml of luciferase buffer. The coelenterazine luciferase assay buffer described in Matthews et al., Bioch. 16 85-91, (1977) incorporated herein by reference is prepared. 50 $\mu$l of pure luciferase (prepared as described in Matthews et al., *supra*) is added to the vial. Corrected emission spectra are collected on an on-line computerized fluorimeter.

The luminometer gives a reading of 1.5 x $10^7$ hv/sec upon addition of luciferase, against a background of 4 x $10^5$ hv/sec. The read-out from this instrument appears in Fig. 1.

## EXAMPLE V

*E. coli* BLR (DE3) is grown in plates under anaerobic conditions in a Gas-Pak container according to the instructions of the manufacturer (Becton Dickinson Microbiology Systems). Colony growth is slowed, due, it is believed, to the anaerobic conditions. The resulting colonies do not detectably exhibit bioluminescence after at least 3 days of growth under anaerobic conditions. However, after being exposed to air for 24 hours, the colonies do begin to exhibit bioluminescence.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: Ward, William
Chalfie, Martin

(ii) TITLE OF INVENTION: BIOLUMINESCENT INDICATOR FOR GENE EXPRESSION AND DETECTION OF MUTAGENESIS BASED UPON THE EXPRESSION OF A GENE FOR A MODIFIED GREEN-FLUORESCENT PROTEIN

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Omri M. Behr, Esq.
(B) STREET: 325 Pierson Avenue
(C) CITY: Edison
(D) STATE: New Jersey
(E) COUNTRY: USA
(F) ZIP: 08837

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/192,158
(B) FILING DATE: 04-FEB-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Behr, Omri M.
(B) REGISTRATION NUMBER: 22,940
(C) REFERENCE/DOCKET NUMBER: RUTG3.0-017

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (908) 494-5240
(B) TELEFAX: (908) 494-0428
(C) TELEX: 51 1642 BEPATEDIN

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 238 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: one-of(1)
(D) OTHER INFORMATION: /note= "Residue 1 Xaa = Methionyl-alanine"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Xaa Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1               5                   10                  15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
            20                  25                  30
```

```
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
        35                  40                  45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
    50                  55                  60

Tyr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
65                  70                  75                  80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                85                  90                  95

Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Ser Arg Ala Glu Val
            100                 105                 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
        115                 120                 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
    130                 135                 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145                 150                 155                 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
                165                 170                 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
            180                 185                 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
        195                 200                 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
    210                 215                 220

Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 944 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
ATGGCTAGCA AAGGAGAAGA ACTTTTCACT GGAGTTGTCC CAATTCTTGT TGAATTAGAT 60
GGTGATGTTA ATGGGCACAA ATTTTCTGTC AGTGGAGAGG GTGAAGGTGA TGCAACATAC 120
GGAAAACTTA CCCTTAAATT TATTTGCACT ACTGGAAAAC TACCTGTTCC ATGGCCAACA 180
CTTGTCACTA CTTTCTATTA TGGTGTTCAA TGCTTTTCAA GATACCCAGA TCATATGAAA 240
CAGCATGACT TTTTCAAGAG TGCCATGCCC GAAGGTTATG TACAGGAAAG AACTATATTT 300
TTCAAAGATG ACGGGAACTA CAAGACACGT GCTGAAGTCA AGTTTGAAGG TGATACCCTT 360
GTTAATAGAA TCGAGTTAAA AGGTATTGAT TTTAAAGAAG ATGGAAACAT TCTTGGACAC 420
```

AAATTGGAAT ACAACTATAA CTCACACAAT GTATACATCA TGGCAGACAA ACAAAAGAAT 480

GGAATCAAAG TTAACTTCAA AATTAGACAC AACATTGAAG ATGGAAGCGT TCAACTAGCA 540

GACCATTATC AACAAAATAC TCCAATTGGC GATGGCCCTG TCCTTTTACC AGACAACCAT 600

TACCTGTCCA CACAATCTGC CCTTTCGAAA GATCCCAACG AAAAGAGAGA CCACATGGTC 660

CTTCTTGAGT TTGTAACAGC TGCTGGGATT ACACATGGCA TGGATGAACT ATACAAATAA 720

ATGTCCAGAC TTCCAATTGA CACTAAAGTG TCCGAACAAT TACTAAAATC TCAGGGTTCC 780

TGGTTAAATT CAGGCTGAGA TATTATTTAT ATATTTATAG ATTCATTAAA ATTGTATGAA 840

TAATTTATTG ATGTTATTGA TAGAGGTTAT TTTCTTATTA AACAGGCTAC TTGGAGTGTA 900

TTCTTAATTC TATATTAATT ACAATTTGAT TTGACTTGCT CAAA 944

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTCTATTATG GTGTTCAA 18

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCTGTTCCAT GGCCAACACT TGTCACTACT TTCTATTATG 40

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Phe Ser Tyr Gly Val Gln
1 5

WE CLAIM:

1. A pre-coelenterazine peptide comprising a modified *Aequorea victoria* GFP wherein $R^{65}$ is Tyr.

2. The pre-coelenterazine peptide of Claim 1 comprising at least amino acid residues $R^{1}$ through $R^{228}$ of said modified *A. victoria* GFP.

3. The pre-coelenterazine peptide of Claim 1 comprising amino acid residues $R^{1}$ through $R^{238}$ of said *A. victoria* GFP, wherein $R^{80}$ is Gln or Arg, $R^{100}$ is Phe or Tyr, $R^{108}$ is Thr or Ser, $R^{141}$ is Leu or Met, $R^{172}$ is Glu or Lys, and $R^{219}$ is Val or Ile.

4. The pre-coelenterazine peptide of Claim 3 wherein $R^{80}$ is Gln, $R^{100}$ is Phe, $R^{108}$ is Thr, $R^{141}$ is Leu, $R^{172}$ is Glu and $R^{219}$ is Val.

5. The pre-coelenterazine peptide of Claim 3 wherein $R^{80}$ is Gln, $R^{100}$ is Tyr, $R^{108}$ is Ser, $R^{141}$ is Met, $R^{172}$ is Glu and $R^{219}$ is Ile.

6. The pre-coelenterazine peptide of Claim 4 wherein $R^{1}$ is methionyl-alanine.

7. A polynucleotide comprising one or more sequences of nucleotide bases collectively encoding the pre-coelenterazine peptide of Claim 1.

8. A polynucleotide comprising one or more sequences of nucleotide bases collectively encoding the pre-coelenterazine peptide of Claim 4.

9. The polynucleotide of Claim 8 comprising the cDNA polynucleotide gfp(C197A).

10. The polynucleotide of Claim 7 further comprising one or more sequences of nucleotide bases collectively encoding the amino acid sequence of a luciferase compatible with coelenterazine.

11. The polynucleotide of Claim 7 further comprising one or more sequences of nucleotide bases collectively encoding the amino acid sequence of apo-aequorin.

12. An expression vector comprising the polynucleotide of Claim 7, and further comprising one or more sequences of nucleotide bases which encode at least one regulatory element operatively linked to said one or more sequences encoding said pre-coelenterazine peptide.

13. An expression vector of Claim 12 further comprising one or more sequences of nucleotide bases which collectively confer resistance to an antibiotic upon an organism transformed therewith.

14. The expression vector of Claim 12 comprising the nucleotide sequence of plasmid TU#132 (ATCC Accession No. 75666).

15. The expression vector of Claim 12 wherein said regulatory element is a promoter selected from the group consisting of promoters from a P450 gene, a promoter activated by a heavy metal, and a promoter from a gene encoding a stress protein.

16. The expression vector of Claim 12 further comprising one or more sequences of nucleotide bases collectively encoding a luciferase compatible with coelenterazine.

17. The expression vector of Claim 16 further comprising one or more sequences of nucleotide bases encoding a further regulatory element

operatively linked to said one or more sequences of nucleotide bases encoding said luciferase.

18. The expression vector of Claim 17 wherein said at least one regulatory element operatively linked to said one or more sequences encoding said pre-coelenterazine peptide is the same as said further regulatory element operatively linked to said one or more sequences encoding luciferase.

19. The expression vector of Claim 17 wherein said at least one regulatory element operatively linked to said one or more sequences encoding said pre-coelenterazine peptide differs from said further regulatory element operatively linked to said one or more sequences encoding luciferase.

20. The expression vector of Claim 12 further comprising one or more sequences of nucleotide bases collectively encoding apo-aequorin.

21. An organism transformed with the polynucleotide of Claim 12.

22. The organism of Claim 21 which is an animal, bacterial, plant or insect cell.

23. The animal cell of Claim 22 which is selected from the group consisting of invertebrate, vertebrate and mammalian cells.

24. An organism transformed with the expression vector of Claim 14.

25. The organism of Claim 24 selected from the group consisting of *E. coli* BLR (DE3) and *E. coli* SMC2 (ATCC Accession No. 69553).

26. An organism of Claim 21 further transformed with a second polynucleotide comprising one or more sequences of nucleotide bases collectively encoding a luciferase compatible with coelenterazine.

27. An organism transformed with the expression vector of Claim 16.

28. An organism of Claim 21 further transformed with a second polynucleotide comprising one or more sequences of nucleotide bases collectively encoding apo-aequorin.

29. An organism transformed with the expression vector of Claim 20.

30. The organism of Claim 29 which is a squid giant neuron.

31. A method of expressing the polynucleotide of Claim 7 comprising incubating said polynucleotide in the presence of means for effecting expression of said polynucleotide under conditions favorable to expression of said polynucleotide.

32. A method of synthesizing coelenterazine comprising expressing said polynucleotide according to Claim 31 and collecting coelenterazine from said means.

33. The method of Claim 32, wherein said step of incubating said polynucleotide is preceded by transforming said organism with said polynucleotide; said means for effecting expression of said polynucleotide is an organism transformed with said polynucleotide; said step of incubating said polynucleotide in the presence of said means comprises culturing said transformed organism for one or more generations under conditions favorable to growth of said transformed organism and favorable to expression of

said polynucleotide; and said step of collecting coelenterazine is performed by lysing the progeny of said cultured transformed organism to form a cell-free extract.

34. The method of Claim 33 wherein said means for effecting expression of said polynucleotide is a cell selected from the group consisting of *E. coli* SMC2 (ATCC Accession No. 69553) and *E. coli* BLR (DE3) transformed with an expression vector comprising said polynucleotide.

35. The method of Claim 33, wherein said organism is cultured aerobically.

36. The method of Claim 33 wherein said transformed organism is cultured in the presence NADP.

37. The method of Claim 31 wherein said polynucleotide is a polyribonucleotide, and said means for effecting expression of said polyribonucleotide is a cell-free aqueous translation system.

38. The method of Claim 30 further comprising converting said collected coelenterazine to luciferyl sulfate.

39. The method of Claim 38, wherein said converting is performed by incubation of said coelenterazine with a luciferin sulfokinase.

40. Purified coelenterazine synthesized by the method of Claim 32.

41. Purified luciferyl sulfate synthesized by the method of Claim 38.

42. A method for selecting cells expressing a protein of interest, wherein said cells comprise a polynucleotide comprising one or more

sequences of nucleotide bases collectively encoding said protein of interest and further comprising a regulatory element operatively linked to said encoding sequences, said method comprising

a) transforming said cells with the expression vector of Claim 10;

b) culturing said cells under conditions permitting expression of said pre-coelenterazine peptide and the protein of interest; and

c) selecting the cultured cells which express coelenterazine, thereby selecting cells expressing the protein of interest.

43. A method for detecting expression of a gene of interest in a cell which comprises:

a) introducing into a cell a polynucleotide comprising one or more sequences of nucleotide bases collectively encoding a regulatory element and said gene of interest, and a polynucleotide of Claim 5, such that the regulatory element of the gene controls expression of pre-coelenterazine peptide;

b) culturing said cell in conditions permitting expression of the gene of interest and of said pre-coelenterazine peptide; and

c) detecting the expression of coelenterazine in the cell, thereby indicating the expression of the gene in the cell.

44. A method for detecting increased levels of intracellular calcium ion, comprising

a) culturing an organism of Claim 28 under conditions favorable to growth of said transformed organism and favorable to expression of said pre-coelenterazine and apo-aequorin peptides; and

b) monitoring said culture for exhibition of bioluminescence.

45. A method for detecting the presence of $O_2$ in an anaerobic system, comprising

a) culturing an organism of Claim 26 under conditions favorable to growth of said transformed organism and favorable to expression of said

pre-coelenterazine and luciferase peptides, where said organism is a facultative anaerobe; and

b) monitoring said culture for exhibition of bioluminescence.

46. The organism of Claim 26 wherein said second polynucleotide comprises, in its one or more sequences of nucleotide bases, a mutation which precludes a bioluminescent interaction between an expression product of said second polynucleotide with an expression product of said first polynucleotide, said mutation being reversible upon exposure to a mutagen to enable a bioluminescent interaction between said expression products.

47. A method of testing the mutagenicity of a chemical compound, comprising:

a) growing a culture of said organism of Claim 41 through one or more generations in a nutrient medium comprising said chemical compound; and

b) measuring the bioluminescence of said culture and comparing said bioluminescence to that from a culture of said organisms of Claim 41 grown in the absence of said chemical compound.

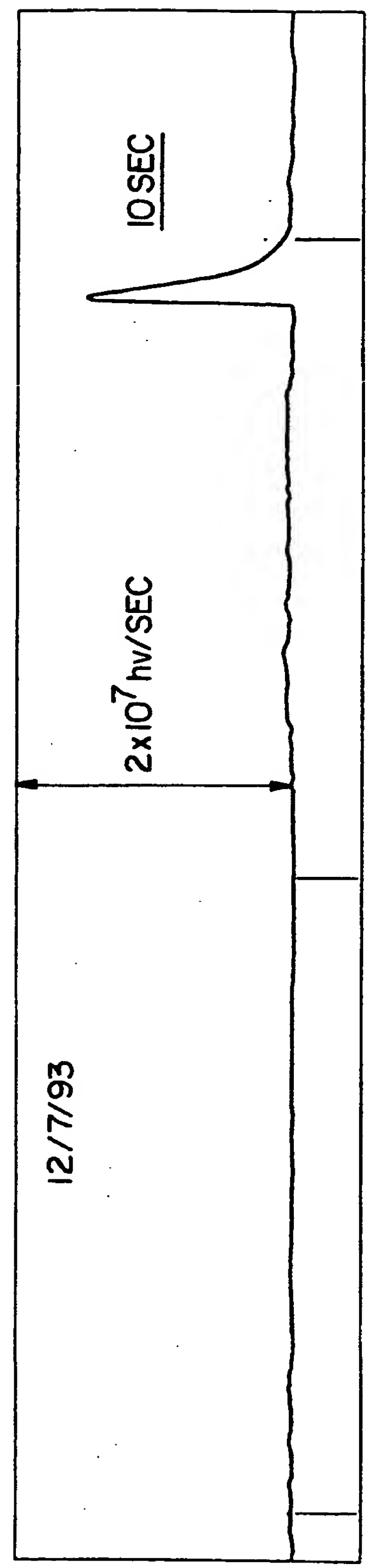



FIG. 1

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/01425

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.2, 23.5; 435/6, 69.1, 119, 240.2, 252.3, 252.33; 544/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y, P | PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, Volume 91, issued December 1994, R. Heim et al., "Wavelength Mutations and Posttranslational Autooxidation of Green Fluorescent Protein", pages 12501-12504, see particularly Figure 2. | 1-47 |
| X<br>---<br>Y | John Burr, "Chemi- and Bioluminescence", published 1985 by Marcel Dekker, Inc., pages 321-386, see particularly, pages 323-329, 336 and 371-376. | 40, 41<br>---------<br>1-39, 42-47 |
| Y, P | SCIENCE, Volume 263, issued 11 February 1994, M. Chalfie et al., "Green Fluorescent Protein as a Marker for Gene Expression", pages 802-805, see entire article. | 1-47 |

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
08 MAY 1995

Date of mailing of the international search report
30 MAY 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
REBECCA PROUTY
Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | BIOCHEMISTRY, Volume 32, No. 5, issued February 1993. C.W. Cody et al., "Chemical Structure of the Hexapeptide Chromophore of the *Aequorea* Green-Fluorescent Protein", pages 1212-1218, see particularly Figure 1. | 1-47 |
| Y, P | FEBS LETTERS, Volume 341, issued March 1994, S. Inouye et al., "*Aequorea* Green Fluorescent Protein. Expression of the Gene and Fluorescence Characteristics of the Recombinant Protein", pages 277-280, see entire article. | 1-47 |
| Y | US, A, 5,162,227 (COMIER) 10 November 1992, see entire document. | 11, 20, 28-30, 44 |
| Y | GENE, Volume 111, issued 1992, D.C. Prasher et al., "Primary Structure of the *Aequorea victoria* Green-Fluorescent Protein", pages 229-233, see entire document. | 1-47 |
| Y | PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, Volume 88, issued May 1991, W.W. Lorenz et al., "Isolation and Expression of a cDNA Encoding *Renilla reniformis* Luciferase", pages 4438-4442, see entire document. | 10, 16-19, 26-27, 45-47 |

## INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/01425

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/01425

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07K 14/435; C12N 1/21, 5/10, 15/09, 15/12, 15/53, 15/63; C12P 17/18; C07D 471/00; C12Q 1/68

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

530/350; 536/23.2, 23.5; 435/6, 69.1, 119, 240.2, 252.3, 252.33; 544/350

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, CA, WPI, MEDLINE, EMBASE, LIFESCI, BIOSIS, BIOTECHDS
search terms: coelenterazine, pre-coelenterazine, coelenterate(2a)luciferin, green fluorescent protein# or gfp, muta? or modif?, gene# or sequence#

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-31 and 46, drawn to a pre-coelenterazine peptide, DNA, vectors and host cells encoding the peptide and method of expressing the DNA to produce the peptide.

Group II, claims 32-41, drawn to coelenterazine and luciferyl sulfate and method of making.

Group III, claim 42, drawn to a cell selection technique.

Group IV, claim 43, drawn to a method of detecting expression of a gene.

Group V, claim 44, drawn to a method of detecting calcium.

Group VI, claim 45, drawn to a method of detecting O2.

Group VII, claim 47, drawn to a method of detecting mutagenicity.

The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The products of Groups I and II are chemically different structures, the products of Group I consisting of peptides and nucleic acids while the products of Group II consist of an imidizolopyrazine and the peptide and nucleic acids of Group I do not incorporate any essential structural element of the product of Group II.

The methods of Groups I and III-VII are all independent and do not share a special technical feature because each method comprises different steps and produces different products or results (i.e, the method of Group I results in the production of a peptide, the method of Group III results in selection of a cell, the method of Group IV results in detection of gene expression, the method of Group V results in detection of calcium, the method of Group VI results in detection of oxygen and the method of Group VII results in detection of mutagenicity).

Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.